

1, 3 . 1, 2, 3 . 2, 3 . 2, 3
1 , 2 , 3

Computer-Aided Diagnosis Using the Digital Database for Screening Mammography

J. H. Lee^{1,3}, H. J. Kim^{1,2,3}, E. K. Kim^{2,3}, K. W. Kim^{2,3}

¹Department of Medical Science, ²Department of Diagnostic Radiology,
³Research Institute of Radiological Science, Yonsei University College of Medicine

= Abstract =

The purpose of this study was to propose robust schemes in terms of mass diagnosis and to objectively describe the confounding effect on cancer risk. For the experiment of mammographic masses, 249 cases of malignant (case group) and 307 cases of benign (control group) from the Digital Database for Screening Mammography (DDSM) were selected. Each radiological feature was categorized and then performed for the multiple logistic regression analysis to reveal the interaction to the probability of a malignant tumor. For 95% confidence interval and high p-value, these analyses significantly showed that the effects on incidence rate of malignant tumor were complexly associated in the order of margins, age, size, shape and breast tissue density. Our method may be useful for identifying breast cancer in mammography and developing computer-aided diagnosis as a solution under the PACS environment.

Key words: Breast cancer, Digital database for screening mammography,
Computer-aided diagnosis, Multiple logistic regression analysis

가 가가 [1].
(mammography)
가
가
2001 (40,191) 16.1% 1 [2, 3].
(mass)가
X
(glandular tissue or parenchyma)
X
가
(: 02 - PJ3 - PG6 - EV08 - 0001). 가 .

Tel: 02 - 361 - 5753, Fax: 02 - 313 - 1039
E - mail: hjkim@yumc.yonsei.ac.kr

가

가

가

가

가

가

2.

South Florida Digital database for Screening
[4, 5]. Mammography(DDSM) [7]

가

2,500

(),

ACR(American College of Radiology)

[6].

가,

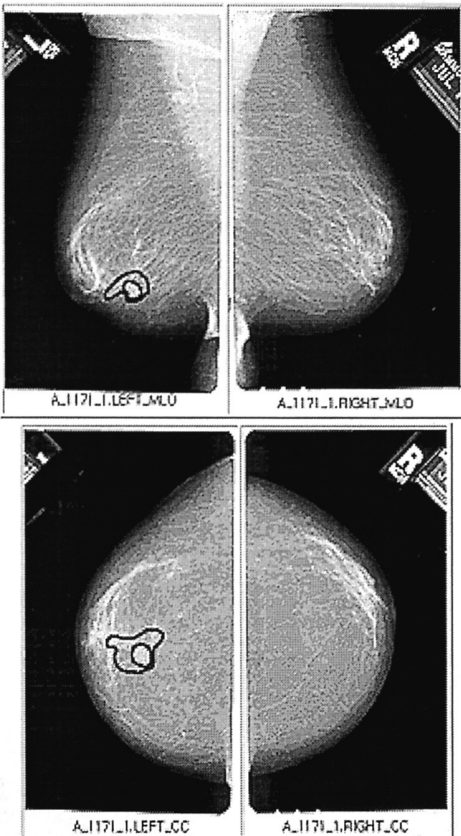
ACR

(,)

가

(ground

Volume: cancer_06 Case: A-1171-1



```

ics_version 1.0
filename A-1171-1
DATE_OF_STUDY 18 7 1994
PATIENT_AGE 73
FILM
FILM_TYPE REGULAR
DENSITY 1
DATE_DIGITIZED 2 7 1998
DIGITIZER HWPTEK 43.5
SEQUENCE
LEFT_CC LINES 5491 PIXELS_PER_LINE 3166 BITS_PER_PIXEL 12 RESOLUTION 43.5 OVERLAY
LEFT_MLO LINES 5401 PIXELS_PER_LINE 3256 BITS_PER_PIXEL 12 RESOLUTION 43.5 OVERLAY
RIGHT_CC LINES 5491 PIXELS_PER_LINE 3001 BITS_PER_PIXEL 12 RESOLUTION 43.5 NON_OVERLAY
RIGHT_MLO LINES 5491 PIXELS_PER_LINE 3421 BITS_PER_PIXEL 12 RESOLUTION 43.5 NON_OVERLAY

```

FILE: A_1171_1.LEFT_MLO.OVERLAY

TOTAL_ABNORMALITIES 1

ABNORMALITY 1

LESION_TYPE MASS SHAPE OVAL MARGINS SPICULATED

ASSESSMENT 5

SUBTLETY 5

PATHOLOGY MALIGNANT

TOTAL_OUTLINES 2

BOUNDARY

CORE

FILE: A_1171_1.LEFT_CC.OVERLAY

TOTAL_ABNORMALITIES 1

ABNORMALITY 1

LESION_TYPE MASS SHAPE OVAL MARGINS SPICULATED

ASSESSMENT 5

SUBTLETY 5

PATHOLOGY MALIGNANT

TOTAL_OUTLINES 2

BOUNDARY

CORE

1. DDSM web page. It is showing the thumbnail version of the images for case 1171. In this case, the patient age is 73 with high breast tissue density and the abnormality is a malignant speculated mass. All cases in DDSM are female patients.

truth) . ACR Breast Imaging
 1990 1998
 249 () 307 () 556
 Reporting and Data System(BI - RADSTM)
 [11].

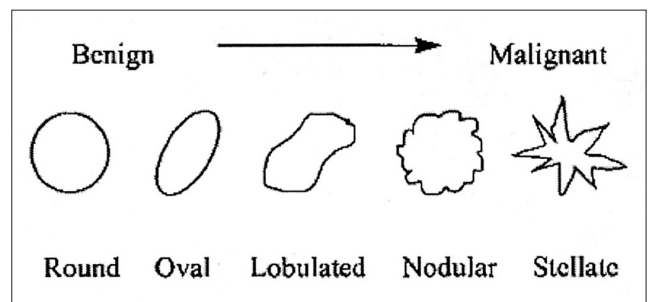
(mediolateral oblique view: MLO) (craniocaudal
 view: CC) 가
 가 (shape)
 1) (round)
 2) (oval)
 3) (lobulated)
 4) (irregular)

totalis) MLO CC 가 (pec -
 (margins)
 1) (circumscribed)
 2) (obscured)
 3) (microlobulated)
 4) 가 (ill - defined)
 5) (spiculated)
 (margin),

(density or intensity)
 (density or attenuation)
 (attenuation)
 [8].
 (density or attenuation)
 1) (high)
 2) (equal or isodense)
 3) (lower attenuation, but not fat containing)
 4) (fat containing or radiolucent)

1 2 2 cm 2.2.
 2cm 가
 가
 [9].
 ROI 가

(distortion) 가 (fibrous projection)
 (spiculated) (radiating)
 가
 가 [10].
 가
 가 (lucent - centered)



2. Morphologic spectrum of mammographic masses.

[10].

(National Institutes of Health, Bethesda, Maryland, USA)

DICOM

X 가

가

Plug -

in

(thresholding)

Segmenting

Assistant Plug - in

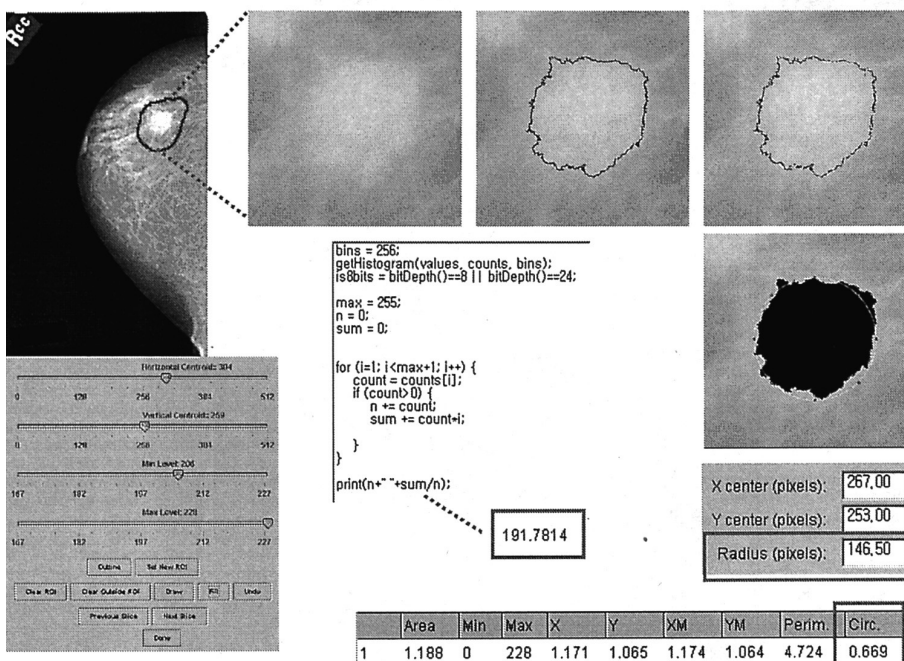
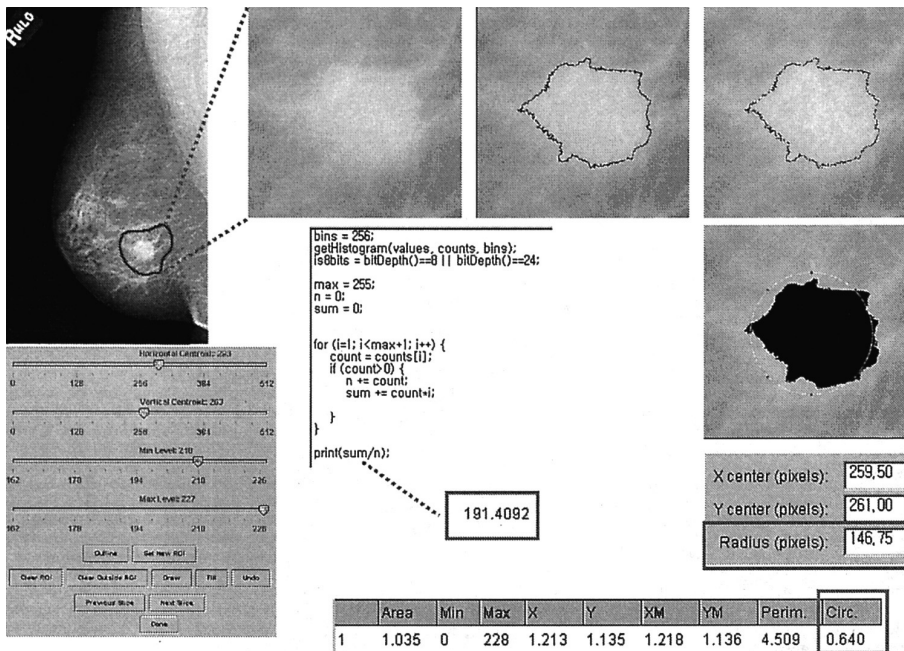
ROI

ROI

DICOM Viewer Image - J ver 1.30

grey level

3. Experimental procedures from the MLO/CC images of an identical patient.



:

가

2.3.

가. : 1.20 , 1.20 - 1.49, 1.50 - 1.79, 1.80

512×512 ROI

. : 1.00cm , 1.00cm - 1.49cm, 1.50cm

가

. : 0.6 , 0.60 - 0.69, 0.70 - 0.79, 0.80

. : 40 , 40 , 50 , 60 , 70

. (cir - : circumscribed/obscured, microlobulated/ ill -
cularity) [12]. defined, spiculated

$Circularity = 4 \cdot (area/perimeter^2)$

area: area of selection in square pixels

가

perimeter: length of the outside boundary of the selec -
tion

SAS System 8.02

ver 8.02

1

,

0

(SAS Institute Inc., Cary, NC, U.S.A.)

가

x, y

MLO CC

0

3.

grey level

3.1.

(normalize)

가

40

5.5%, 40 가 25.9%,

X

step -

50 가 32.3%, 60 가 61.6%, 70

74.1%

wedge

grey level

가

(optical density)

(1).

(Wald)

(Likelihood Ratio)

(Score)

(generalized linear model:

GLM)

가 $H_0: = 0($)

2.4.

가

가

0

가

가

가
가
가
가
1
120.655(p -value 0.0001) H_0 :
=0
0
111.168 (p -value
0.0001), 95.005(p -value 0.0001)
(2).

$z =$
/ASE(asymptotic standard error)=0.836/0.086=9.7
 $z^2=95.0$ (p -value 0.0001)
가 5%
= - 3.023($SE=0.311$),
=0.836($SE=0.086$) 가
 $P(x)$ 가 (3). 가
odds exp()=exp(0.836)=2.307 , 230%
가 (odds ratio) (4).
가 95%
(1.950, 2.729)
95%
grey level

가

1. Table of Age vs. Mammary Tumor

		52	3	55
40	%	94.5%	5.5%	100.0%
40	%	83	29	112
	%	74.1%	25.9%	100.0%
50	%	86	41	127
	%	67.7%	32.3%	100.0%
60	%	56	90	146
	%	38.4%	61.6%	100.0%
70	%	30	86	116
	%	25.9%	74.1%	100.0%
	%	307	249	556
	%	55.2%	44.8%	100.0%

1.80 가 39.2%, 1.50 1.79
가 40.6%, 1.20 1.49
가 51.0%, 1.20
54.2%
(가 가) 가 (5).
1 , ,
 p -value 0.01 =0
가 (6).
= - 0.759($SE=0.226$)
=0.237($SE=0.090$) (가 가) $P(x)$ 가 (7).
1.268 가 가
1.268 (126.8% 가)가 (8).
1.00cm 35.5%, 1.00cm
1.49cm 45.0%, 1.50cm 51.0% (9).
1 (15.207, p -value 0.0001), (15.109, p -value=0.0001), (14.855, p -value=0.0001)
=0 가 (10).
=0.448($SE=0.116$)
가 $P(x)$ 가 가 (11).

x $x+1$ $P/(1 -$

2. Testing Global Null Hypothesis: BETA=0 (age)

Likelihood Ratio	1	120.655	< 0.0001
Score	1	111.168	< 0.0001
Wald	1	95.005	< 0.0001

3. Analysis of Maximum Likelihood Estimation (age)

Parameter	DE	Estimate	Standard Error	Wald Chi - Square	Pr > ChiSq
Intercept	1	- 3.023	0.311	94.447	< 0.0001
Age	1	0.836	0.086	95.005	< 0.0001

4. Odds Ration Estimates (age)

Effect	Point Estimate	95% Wald	Confidence Limits
Age	2,307	1,950	2,729

$P)=1.566$ (12), $\exp()=\exp(0.448)$
 .
 .
 156.6%가 가 .
 0.6
 가 74.4% 가
 0.60 0.69 가 55.8%, 0.70
 0.79 41.9%, 가 가 0.8
 14.6% (13).
 , , ,
 (p -value 0.0001),
 $=0$ 가 (14).
 $=0.880(SE =0.092)$
 ()

5. Table of Density vs. Mammary Tumor

	76	49	125
1.80	60.8%	39.2%	100.0%
1.50 - 1.79	124	84	208
	59.6%	40.6%	100.0%
1.20 - 1.49	74	77	151
	49.0%	51.0%	100.0%
1.20	33	39	72
	45.8%	54.2%	100.0%
	307	249	556
	55.2%	44.8%	100.0%

6. Testing global Null Hypothesis: BETA=0 (density)

Likelihood Ratio	1	7.064	0.0079
Score	1	7.045	0.0079
Wald	1	6.982	0.0082

7. Analysis of Maximum Likelihood Estimation (density)

Parameter	DE	Estimate	Standard Error	Wald Chi - Square	Pr > ChiSq
Intercept	1	- 0.759	0.226	11.299	0.0008
Density	1	0.237	0.090	6.982	0.0082

4. Odds Ratio Estimates (density)

Effect	Point Estimate	95% Wald	Confidence Limits
Density	1,268	1,063	1,512

:

$P(x)$ 가 가 가 (15).
 2.410
 가 가
 2.41 (240% 가) (16).
 circumscribed/obscured
 16.4%, microlobulated/ill - defined
 가 55.2%, (spiculated) 95.7%
 (17).
 $=0$ 가
 (18), 2.121($SE =$
 0.171) (19).

$$P(x=1)=0.150$$

$$P(x=3)=0.925$$

가

9. Table of Mass Size vs. Mammary Tumor

	120	66	186
1.00 cm	64.5%	35.5%	100.0%
1.00 - 1.49 cm	131	107	238
	55.0%	45.0%	100.0%
1.50 cm	56	76	132
	49.0%	51.0%	100.0%
	307	249	556
	55.2%	44.8%	100.0%

10. Testing Global Null Hypothesis: BETA=0 (size)

Likelihood Ratio	1	15.207	< 0.0001
Score	1	15.109	0.0001
Wald	1	14.855	0.0001

11. Analysis of Maximum Likelihood Estimation (size)

Parameter	DE	Estimate	Standard Error	Wald Chi - Square	Pr > ChiSq
Intercept	1	- 1.068	0.240	19.781	< 0.0001
Size	1	0.448	0.116	14.855	0.0001

12. Odds Ratio Estimates (size)

Effect	Point Estimate	95% Wald	Confidence Limits
Size	1,566	1,246	1,966

8.335

가 (20).

3.2.

가

가

p - value 0.0001

p - value=0.0002, p - value=0.0304

(22).

p - value=0.0908

5%

13. Table of Circularity vs. Mammary Tumor

		129	22	151
0.80		85.4%	14.6%	100.0%
		93	67	160
0.70 - 0.79		58.1%	41.9%	100.0%
		53	67	120
0.60 - 0.69		44.2%	55.8%	100.0%
		32	93	125
0.60		25.6%	74.4%	100.0%
		307	249	556
		55.2%	44.8%	100.0%

14. Testing Global Null Hypothesis: BETA=0 (shape)

Likelihood Ratio	1	110.079	< 0.0001
Score	1	104.348	< 0.0001
Wald	1	91.966	< 0.0001

15. Analysis of Maximum Likelihood Estimation (shape)

Parameter	DE	Estimate	Standard Error	Wald Chi - Square	Pr > ChiSq
Intercept	1	- 2.352	0.245	92.165	< 0.0001
Circularity	1	0.880	0.092	91.966	< 0.0001

16. Odds Ratio Estimates (shape)

Effect	Point Estimate	95% Wald	Confidence Limits
Circularity	2,410	2,013	2,885

Y= 1

$$\text{logit} [P (Y =1)]=\log \left[\frac{P (Y =1)}{1 - P (Y =1)} \right]$$

$$= - 8.495+0.817X_1+0.22X_2+0.658X_3+0.263X_4+1.833X_5$$

$$P (Y =1)=$$

$$\frac{\exp(- 8.495+0.817X_1+0.22X_2+0.658X_3+0.263X_4+1.833X_5)}{1+\exp(- 8.495+0.817X_1+0.22X_2+0.658X_3+0.263X_4+1.833X_5)}$$

가

가

$$\text{가 } 2.264=\exp(0.817)$$

가

가

2.264

가

17. Table of Margin vs. Mammary Tumor

		225	44	269
Circumscribed/ Obscured		83.6%	16.4%	100.0%
		77	95	172
Microlobulated /Ill - defined		44.8%	55.2%	100.0%
		5	110	115
Spiculated		4.3%	95.7%	100.0%
		307	249	556
		55.2%	44.8%	100.0%

18. Testing Global Null Hypothesis: BETA=0 (margin)

Likelihood Ratio	1	243.680	< 0.0001
Score	1	215.813	< 0.0001
Wald	1	153.203	< 0.0001

19. Analysis of Maximum Likelihood Estimation (margin)

Parameter	DE	Estimate	Standard Error	Wald Chi - Square	Pr > ChiSq
Intercept	1	- 3.856	0.309	155.815	< 0.0001
Margin	1	2.121	0.171	153.203	< 0.0001

20. Odds Ratio Estimates (margin)

Effect	Point Estimate	95% Wald	Confidence Limits
Margin	8,335	5,958	11,662

(1.814, 2.827) .
 , 1.246 , 1.931 ,
 1.301 ,
 $\exp(1.833)=6.253$ 가 (23).
 , 가
 , (6.253),
 (2.264), (1.931), (1.301),
 (1.246)
 .
 24
 (percent concordant) 90.4%, (percent dis -
 cordant) 9.4%, (percent tied)
 0.2% .
 4 ROC(receiver operating characteris -
 tic)
 (sensitivity) (specificity) .
 , 가
 (90.4%), 556
 가 (AUC: the

area under the curve=0.905).

4.

가 가
 ,
 가
 (double reading)
 가
 screening program .
 .
 (PACS)
 가 . 가 가
 PACS

21. Testing Global Null Hypothesis: BETA=0

Likelihood Ratio	5	329.963	< 0.0001
Score	5	267.373	< 0.0001
Wald	5	157.462	< 0.0001

22. Analysis of Maximum Likelihood Estimation

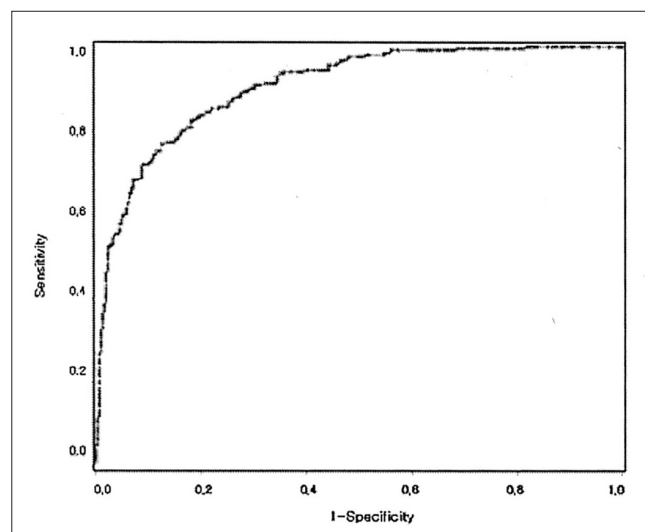
Parameter	DE	Estimate	Standard Error	Wald Chi - Square	Pr > ChiSq
Intercept	1	- 0.495	0.774	120.571	< 0.0001
Age	1	0.817	0.113	52.080	< 0.0001
Density	1	0.220	0.130	2.860	0.0908
Size	1	0.658	0.178	13.700	0.0002
Shape	1	0.263	0.122	4.687	0.0304
Margin	1	1.833	0.201	83.590	< 0.0001

23. Odds Ratio Estimates

Effect	Point Estimate	95% Wald	Confidence Limits
Age	2.264	1.814	2.827
Density	1.246	0.966	1.609
Size	1.931	1.363	2.737
Shape	1.301	1.025	1.650
Margin	6.253	4.221	9.263

24. Association of Predicted Probabilities and Observed Responses

Percent Concordant	90.4%	Somers D	0.809
Percent Discordant	9.4%	Gamma	0.811
Percent Tied	0.2%	Tau - a	0.401
Pairs	76.443	c	0.905



4. ROC curve at a confidence level of 95%.

가

1. . 2003 가 ; 2003
2. . ; 1995. p.57-84
3. Kopans DB. Breast Imaging. Philadelphia: JB Lippincott; 1989. p.1-16
4. Doi K, MacMahon H, Katsuragawa S, Nishikawa RM, Jiang Y. Computer-Aided Diagnosis in Radiology: Potential and Pitfalls. Eur J Radiol 1999;31:97-109
5. Health M, bowyer K, Kopans D, Moore R, Kegelmeyer P Jr. Mass Detection by Relative Image Intensity. In: the Proceedings of the 5th International Workshop on Digital Mammography. Wisconsin: Medical Physics Publishing; 2000. p.46

6. Doi K, Giger ML, Nishikawa RM, Hoffmann KR, MacMahon H, Schmidt RA, et al. Digital Radiography. a Useful Clinical Tool for Computer-Aided Diagnosis by Quantitative Analysis of Radiographic Images. Acta Radiol 1993;34:426-39
7. University of South Florida. DDSM: Digital Database for Screening Mammography. <http://marathon.csee.usf.edu/mammography/Database.html>
8. Malich A, Marx C, Facius M, boehm T, Fleck M, Kaiser WA. Tumour Detection Rate of a New Commercially Available Computer-Aided Detection System. Eur J Radiol 2001;11:2454-9
9. Rangarai M Rangayyan, Nema M El-Faramawy, JE Leo Desautels, Onsy A Alim. Measures of Acutance and Shape for Classification of Breast Tumors. IEEE Trans on Medical Imaging 1997;16(6):799-810
10. Zhimin Huo, Maryellen L Giger, Carl J Vyborny, Funmi I Olopade, Dulcy E Wolverton. Computer-Aided Diagnosis: Analysis of Mammographic Parenchymal Patterns and Classification of Masses on Digitized Mammograms. Proceed-ings of the 20th International Conference IEEE/EMBS 1998;20(2): 1017-20
11. American College of Radiology. Breast Imaging Reporting and Data System(BI-RADS). 3rd ed. 1998
12. John C Russ. The Image Processing Handbook. Florida: CRC Press 1995. p.523-6

= =

South Florida

DDSM

249

가

307

X

가

가 가

, 95%

가

가

가